Liquid–vapor coexistence by molecular dynamics simulation

András Baranyai(a)
Department of Theoretical Chemistry, Eötvös University, 1518 Budapest 112, P. O. Box 32, Hungary

Peter T. Cummings
Department of Chemical Engineering, University of Tennessee, 419 Dougherty Engineering Building, Knoxville, Tennessee 37996-220, and Chemical Technology Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6268

(Received 20 September 1999; accepted 3 December 1999)

We present a simple and consistent molecular dynamics algorithm for determining the equilibrium properties of a bulk liquid and its coexisting vapor phase. The simulation follows the dynamics of the two systems simultaneously while maintaining the volume and the number of particles of the composite system fixed. The thermostat can constrain either the total energy or the temperature at a desired value. Division of the extensive properties between the two phases is governed by the difference of the corresponding intensive state variables. Particle numbers are continuous variables and vary only in virtual sense, i.e., the real sizes of the two systems are the same and do not change during the course of the simulation. Calculation of the chemical potential is separate from the dynamics; thus, one can replace the particle exchange step with other method if it improves the efficiency of the code. © 2000 American Institute of Physics. [S0021-9606(00)50608-4]

I. INTRODUCTION

In a previous paper1 we presented a molecular dynamics version of the Gibbs ensemble Monte Carlo (GEMC) algorithm.2,3 The method used deterministic equations of motion with properly chosen constraints representing the macroscopic conditions of thermal and mechanical equilibrium. However, in the case of the equilibrium condition for chemical potentials we applied the stochastic particle exchange method of the GEMC scheme. Thus, our algorithm preserved its simplicity in the equations of motion but became a hybrid approach relying on the assumed ergodicity and mixing of the dynamics. An important shortcoming of the method was its numerical sensitivity to the insertion step which required extra care preventing the destabilization of the integrator.

The algorithm presented in the following is an improved, more consistent version of the one described above1,4 because it avoids the problematic insertion step. Particle numbers in the subsystems do not change during the course of the simulation because the constraint relations do not connect the real subsystems but their virtual counterparts. Extensive properties of the real system are multiplied by a weight function providing the extensive property of the virtual system. The role of a weight function is played by the virtual number of particles which is a continuous variable. The chemical potential as an input parameter determined independently from the dynamics governs the change of these virtual particle numbers. In this sense, the algorithm is more general than the Gibbs ensemble Monte Carlo method because it is not connected to the particle exchange step; any other suitable chemical potential calculation procedure, including thermodynamic integration, can serve this purpose.

In Sec. II we present the algorithm. We repeat the most important parts of our earlier derivation to help the reader in understanding the new developments. In Sec. III we discuss the numerical properties of the algorithm based on some model calculations. In Sec. IV we conclude this study.

II. DESCRIPTION OF THE METHOD

Conditions of microcanonical constraints for the two simulation cells, each chosen from regions deep inside phase 1 and phase 2 are the following:

\[ V = V_1 + V_2, \quad N = N_1 + N_2, \quad E = E_1 + E_2, \]  

(1)

where the subscript indices refer to the corresponding phases. Conditions of equilibrium between the two phases can be expressed by the equivalence of he intensive parameters conjugate to the extensive state variables of Eq. (1),

\[ P_1 = P_2, \quad \mu_1 = \mu_2, \quad T_1 = T_2. \]  

(2)

The most suitable equations of motion for the microscopic dynamics are the following:1,4,6

\[ \dot{q}_i = \frac{p_i}{m_i} + \epsilon(q_i - R_0), \]

\[ \dot{p}_i = F_i - \epsilon p_i - \alpha p_i, \]  

(3)

where \( q_i \) and \( p_i \) are the position and momentum of particle \( i \), \( F_i \) is the force acting on it, \( m_i \) is its mass, and \( R_0 \) is the center of mass of the system. The variable \( \epsilon \) acts as a strain rate factor coupling the system to a barostat and \( \alpha \) is the feedback multiplier coupling the system to a heat bath.

Using \( \epsilon \) we can express the change in the volume

\[ \dot{V}_j = 3V_j \epsilon_j, \]

or

\[ \dot{L}_j = L_j \epsilon_j, \]  

(4)

(a)Electronic mail: baranyai@para.chem.elte.hu
where $L_j$ is the length of cubic box, $j=1,2$. In order to obtain equations for relating $\dot{\epsilon}$ while maintaining the restriction of constant volume we have to differentiate the first condition of Eq. (1) twice with respect to time,

$$V_1(3\epsilon_1^2+\dot{\epsilon}_1)+V_2(3\epsilon_2^2+\dot{\epsilon}_2)=0.$$  (5)

Then the expressions for $\dot{\epsilon}$-s are the following:  

$$\dot{\epsilon}_1=\frac{1}{V_1Q_\epsilon}(P_1-P_2)-\frac{3}{2}(\epsilon_1^2+\epsilon_2^2V_2/V_1),$$  (6a)

$$\dot{\epsilon}_2=\frac{1}{V_2Q_\epsilon}(P_2-P_1)-\frac{3}{2}(\epsilon_2^2+\epsilon_1^2V_1/V_2),$$  (6b)

where $Q_\epsilon$ is the fictitious mass of the barostat. So far our derivation is identical to that of Ref. 4. Now, let us introduce a new coupling parameter, $\xi$, which couples the number of particles to the chemical potential condition. At first, this idea seems awkward, since $N$ is an integer variable not suitable to the continuous nature of differential equations. In addition to this, the number of particles determine the number of equations of motion but have no direct influence on Eq. (3) itself. Still, it is surprising how easy it is to circumvent these obstacles. Let us make $N_1$ and $N_2$ continuous variables and at start let $N_1(t=0)=N/2$ and $N_2(t=0)=N/2$. The actual numbers of particles in the boxes are $N/2$ and $N/2$ and we are not going to change them during the course of the simulation. However, let $N_1(t)$ and $N_2(t)$ refer to the virtual particle numbers of the systems at time $t$ with the meaning that at this time $2N_1/N$ and $2N_2/N$ parts of the simulation cells are considered in the constraints. They can adopt any number between 0 and $N$ continuously but their sum is strictly fixed at $N$.

As a result of the following equations, the relative weights of the two systems will be determined by the chemical potential equilibrium. We can copy the derivation of Eq. (6) by assuming the relationship for phase $j, j=1,2$:

$$\dot{N}_j=3N_\xi \dot{\xi}_j.$$  (7)

Then

$$N_1(3\xi_1^2+\dot{\xi}_1)+N_2(3\xi_2^2+\dot{\xi}_2)=0.$$  (8)

This leads to equations in complete analogy with Eq. (6),

$$\dot{\xi}_1=\frac{1}{N_1Q_\xi}(\mu_2-\mu_1)-\frac{3}{2}(\xi_1^2+\xi_2^2N_2/N_1),$$  (9a)

$$\dot{\xi}_2=\frac{1}{N_2Q_\xi}(\mu_1-\mu_2)-\frac{3}{2}(\xi_2^2+\xi_1^2N_1/N_2),$$  (9b)

where $\mu=\ln(N/V)+\mu^\text{ex}$. The superscript $\text{ex}$ (excess) refers to the difference between the total chemical potential and the perfect gas value at the same temperature and number density. Note that the chemical potential difference in Eq. (9) is the other way around because the phase with smaller chemical potential is the one which increases.

The density of a phase can change for two reasons. Pressure differences can redistribute the volume or chemical potential differences can redistribute the number of particles. [See the differences of $V_j \dot{\epsilon}_1-V_2 \dot{\epsilon}_2$ using Eq. (6) or $N_j \dot{\xi}_1-\dot{N}_j \dot{\xi}_2$ using Eq. (9).] The volumes at start [$V_1(t=0)+V_2(t=0)=V$] are the real volumes of the subsystems with number densities of $\rho_1(t=0)=N/2V_1(t=0)$ and $\rho_2(t=0)=N/2V_2(t=0)$. It is important to note, however, that afterwards $V_1(t)$ and $V_2(t)$ are no longer the real volumes of the real systems with equal numbers of particles but the volumes of the virtual $N_1(t)$ and $N_2(t)$ particle systems. Clearly, the number density of the real system and its virtual counterpart is always identical.

To make the whole idea easier to understand let us imagine the two boxes of the simulation each containing 108 particles. Box 1 represents the bulk phase of the liquid, while box 2 mimics the bulk phase of the vapor. The equality of their pressures and temperatures is not sufficient to ensure the intended separation of the two phases. It is possible that at least one of the boxes contains liquid and vapor phases simultaneously. The essence of the GEMC scheme is to identify the interface particles by their higher chemical potential and move them to the pure phases. Let us say that at a given $(T,V,N)$ condition the composite system is so equilibrated that box 1 and box 2 contain 90 and 126 particles, respectively. To reach this equilibrium state the GEMC algorithm has to transfer 18 particles from system 1 to system 2. If we performed the same simulation using our method the virtual particle numbers in the boxes would be identically 90 and 126. Since, in reality, both of our boxes contain 108 particles, it is the different weight of the systems which ensures the correct division between the two phases. The extensive properties (volume, energy) of the real box 1 take part in the constraints with a weight of 90/108, while this number for the other system is 128/108. In the GEMC scheme the sum of the two system volumes ($V_1, V_2$) is fixed. In our method it is the sum of the weighted volumes which is fixed. Clearly, these weighted volumes are identical to the volumes ($V_1, V_2$) obtained by the GEMC method. The sum of our real volumes ($V'$) corresponds to a different boundary condition $(T,V',N)$ which can be established with subsystems of equal numbers of particles at equilibrium.

In the constrained algorithm of two subsystems, we have some freedom to choose which one of the two conjugate variables is to be fixed and how, since the constraint can be either differential or integral. Constraining the chemical potential in a deterministic algorithm is practically impossible no matter which type of constraint is used. The pressures can be tied to a constant external pressure by an integral feedback but this, in fact, leads to separate systems $[(T,P,N)$ system or isotherm–isobar ensemble] or, in the case of fixed total energy, a fairly unusual composite system $(E,P,N)$. Since it is true in general that the integral feedback gives better numerical accuracy, we did not study the possibility of the volume and number constraints in differential form.

However, in the case of thermal equilibrium we have all the options. The original thermodynamic boundaries of Eq. (1) describe the composite system as a member of the microcanonical ensemble. The total energy can be fixed both ways. The integral feedback leads to fairly complicated expressions because the second derivatives of the forces are also involved. Therefore, we do not present this derivation, but show the simpler differential feedback case.
the ideas of the derivation are given in Ref. 4, we emphasize only the differences due to the new handling of particle numbers. The constraint condition of Eq. (1) for the energy is

\[ \dot{E}_1 + \dot{E}_2 = N_1 e_1 + N_1 \dot{e}_1 + N_2 e_2 + N_2 \dot{e}_2 = 0, \tag{10} \]

where \( e_1 \) is the average energy per particle in phase 1. The first equation for \( \alpha_1 \) and \( \alpha_2 \) follows from Eq. (9) in Ref. 4. The second equation can be created requiring identical starting temperatures, \( T_1(0) = T_2(0) \) and maintaining the \( T_1(t) = T_2(t) \) equality for any \( t \). Then the final formula for \( \alpha_1 \) is

\[
\alpha_1 = \left[ 3(E_1 \xi_1 - V_1 p_1 \epsilon_1) - N_1 e_1 C / \nu_1 + E_2 \xi_2 - V_2 p_2 \epsilon_2 - N_2 e_2 C / \nu_2 + \frac{2N_2}{N} \left( \sum_i F_i p_i / m \right)_1 \right. \\
\left. - \left( \sum_i F_i p_i / m \right)_2 - \epsilon_1 K_1 + \epsilon_2 K_2 \right] / 2K_1, \tag{11} \]

where \( p_j \) is the pressure without long-range correction, \( C \) is a cutoff dependent factor in the long-range correction of the energy [for Lennard-Jones particles \( C = 8 \pi \rho r_c^5 / 3 - r_c^{3} / (\nu) \)], \( \nu_j = V_j / N_j \), and \( K_j = \sum_i p_i^2 / m \). Swapping the corresponding indices in Eq. (11) we obtain the expression for \( \alpha_2 \). Being a differential feedback, this energy constraint will be less accurate which requires regular momentum adjustments much more frequently than in the cases of the volume or number of particles.

If the composite system is a member of the canonical ensemble we maintain a constant temperature in both phases. This can be the Nosé–Hoover form (integral feedback) \(^7\)

\[ \dot{\alpha} = v^2 (T/T_0 - 1), \tag{12} \]

or the Gaussian form (differential feedback) \(^8\)

\[ \alpha_j = \left( \sum_i F_i p_i \right)_j / \left( \sum_i p_i^2 \right)_j - \epsilon_j. \tag{13} \]

We prefer the latter because the kinetic temperatures are identical in the two phases and a synchronous fluctuations in them around the \( T_0 \) target temperature do not influence the dynamics.

The choice of Eqs. (3), (4), (6), (7), (9), and (13) represents the canonical ensemble, the usual GEMC boundary conditions for the composite system. In Sec. III we will present our model calculations based on these equations.

Before passing to the discussion of the practical details of the algorithm, it is appropriate to place our method into the general picture of liquid–vapor phase equilibrium simulation techniques. Chemical potentials, as with other entropic quantities, are functionals of the phase space variables. There are numerous studies in the literature dealing with the problem of direct chemical potential calculations. \(^9,10\) It is outside the scope of the present paper to review all the different possibilities. Clearly, the proliferation of these methods is caused primarily by the need of computationally efficient and reliable algorithms exploiting special possibilities in different models. Since our equations of motion use the chemical potential as an input parameter, the choice of chemical potential calculation can be any of these methods. A generally applicable approach is Kirkwood’s, \(^11\) which was freed from its numerical problems recently by Beutler et al. \(^12\)

Nevertheless, whatever method we choose for direct chemical potential determination, this will take the major share of the computer time. Thus, for complex model systems it is economic to combine our technique with the Gibbs–Duhem integration of Kofke \(^13\) which avoids direct estimations of the chemical potential. The latter method accurately traces coexistence lines of the phase diagram provided variables of its starting state are already on the coexistence line. However, to locate a point on the coexistence line, one cannot avoid determining entropic quantities directly. For the determination of the starting point our method is computationally efficient and generally applicable.

### III. RESULTS AND DISCUSSION

As a test case we chose the exhaustively studied problem of liquid–vapor equilibrium in the Lennard-Jones fluid. At start we placed 108 particles in each cubic box, applied the fifth order Gear integrator and half of the usual reduced time step (0.002) for greater accuracy. (Results are reported in the usual reduced units of computer simulations. \(^14\))

For the sake of simplicity for chemical potential calculation we applied the test particle method. \(^9,10\) At each time step we placed 2000 particles randomly into the cubic boxes. After 500 time steps we averaged the results and calculated the chemical potentials. So each chemical potential value was based on \( 5 \times 10^5 \) test insertions. However, we emphasize that this choice was merely a matter of convenience and any other method could have been used. We reproduced the liquid–vapor results of Ref. 3 with comparable accuracy. Thus, we are not going to report them. Instead we focus our attention on the basic properties of the algorithm.

First, we present some results to show the convergence of the state variables toward their correct equilibrium value. The target equilibrium was a Lennard-Jones liquid–vapor coexistence at \( T = 1.0 \). \(^3\) The chosen starting densities were \( \rho_1 = 0.8 \) and \( \rho_2 = 0.028 \). While the vapor density is within the uncertainties of the target value, the liquid density is higher by about 0.1. We equilibrated the two systems separately not switching on the pressure and chemical potential coupling. The energies \( E_1 = -4.05 \) and \( E_2 = 1.25 \), and the pressures \( P_1 = 0.972 \) and \( P_2 = 0.0242 \) were very close to their target values in the case of the vapor phase (phase 2). First, we switched off the chemical potential coupling and studied only the pressure by using three different fictitious masses in Eq. (5). In Figs. 1, 2, and 3 we show the change of the density and the pressure of the liquid phase when \( Q_s \) is set to 5.0, 1.0, and 0.2, respectively. Data points of the figures represent averages over units of the reduced time (500 time steps). Due to its much larger volume, the vapor phase did not change too much during the adjustment of the liquid phase. The liquid pressure had to move toward the much more stable vapor pressure which on the scales of the figures is hardly distinguishable from zero. The speed toward equilibrium were determined by the value of the fictitious mass. In the case of the smallest mass, it took only 2 time units to get to the range of the vapor pressure. It can be seen in the
figures that the finite memory of the differential equations produced large oscillations around the target density. The size of the amplitude of these oscillations is determined by the starting value, while the frequency of it is determined by the fictitious mass. Although these oscillations are dampening with time, they have no physical relevance and being a computational nuisance can be eliminated from the calculations.

While the finite memory of the differential equation can make the convergence very fast, the price paid is the subsequent long oscillations of the related extensive property. Since the GEMC method simulates a Markov process, only the instantaneous state of the system has an influence on the subsequent state. Nevertheless, we can terminate the memory of the system quite easily and without a direct human intervention. The algorithm can monitor the pressure and chemical potential differences of the coexisting systems. If the differences are smaller than a preset threshold the algorithm may set the derivatives of \( e, L, \) and \( \xi, N \) to zero. This will slightly upset the accuracy of the constraints but with appropriate rescaling the dynamics is directed to a better track.

The second set of experiments started from the same state as the previous but this time we switched off the pressure coupling and equilibrated only with the chemical potentials. In Figs. 4, 5, and 6 these results are shown with \( Q_\xi \) being 10.0, 2.0, and 0.4, respectively. In these figures instead of the liquid density the fictitious number of particles in the liquid phase and the liquid and vapor chemical potentials are shown. Since the total number of particles in the system was fixed at 216, the population of both phases can be read from these figures. The increasing number of particles in the vapor phase increases the chemical potential of that phase because the main contribution comes from the density term. While estimation of the vapor phase chemical potential is quite accurate, there are large uncertainties in the estimated value of the liquid chemical potential. The difference of these two chemical potentials determine the change in the number of particles. Chemical potentials are much more sensitive quantities than the pressure. If one uses too small fictitious mass in the constraint, the dynamics might get out of control (see Fig. 6).

FIG. 1. The density (closed circles) and the pressure (open squares) of the liquid phase for 20 time units after starting the simulations with \( Q_x = 5.0 \). See the text for further details.

FIG. 2. The same as Fig. 1 with a different fictitious mass for the dynamics of the volume. \( (Q_x = 1.0) \)

FIG. 3. The same as Fig. 1 with a different fictitious mass for the dynamics of the volume. \( (Q_x = 0.2) \)

FIG. 4. Chemical potential of the liquid (open squares) and the vapor (closed squares) phase and the number of particles in the liquid phase (closed circles) for 20 time units after starting the simulation with \( Q_\xi = 10.0 \). See the text for further details.
Since in this experiment we did not break up the memory of the system, the generated oscillation went up to very high density which produced unrealistically high chemical potential for the dense phase. Although we did not apply it in these calculations, the best strategy to determine the chemical potential is to create a histogram of this quantity in terms of the density. Empty bins of the histogram could be filled up by interpolated values. As the simulation progresses the histogram can be refined by additional chemical potential estimations. In this scheme, one may use small fictitious masses even for this quantity because at each instant the chemical potential of the phases is identified.

In the following, we present some figures representing the fluctuations in the system when we started it from close to equilibrium position ($r_l = 0.7$ and $r_v = 0.028$) for the same temperature ($T = 1.0$). As can be seen from Fig. 7 the liquid density has several oscillations around its equilibrium value which is a combined result of the simultaneous change of its volume and number of particles (see Fig. 8). Since the mass of the barostat was only 0.5, changes in the volume followed quickly the much slower changes in particle numbers ($Q_j = 20.0$). Redistribution of particles played the crucial role in determining the vapor density because of the large difference of the phase volumes. (Compare Figs. 7 and 8.) The need for comparable masses for pressure and chemical potential feedbacks is evident from these figures. The slow changes in the numbers of particles produced only a half wave during 100 reduced time units. Figure 9 shows the synchronization between the density and the chemical potential of the liquid phase while Figs. 10 and 11 compare the behavior of the intensive quantities for the two phases.

In addition to the canonical composite system, we performed some calculations for the microcanonical composite system as well. Using the test particle method for chemical potential determination, it is a disadvantage of the energy constraint that we do not know the final equilibrium temperature a priori. Since the instantaneous temperature in a small system can have large fluctuations, the chemical potential is estimated at very different temperatures which gives, in general, lower average chemical potential values than the constant temperature estimations.
IV. CONCLUSIONS

We proposed a simple and easy to implement algorithm for simulating liquid–vapor equilibrium by molecular dynamics. Similar to the Gibbs ensemble Monte Carlo method the algorithm mimics the properties of bulk phases coupled by the requirement of constant total volume and number of particles. In the GEMC scheme the temperature is fixed. In our case either the temperature or the total internal energy can be constrained.

We performed representative calculations for Lennard-Jones particles to study the performance of the algorithm. Our method has several advantages over the GEMC method.

1. For complicated molecular models it is generally simpler to devise a MD code than a MC one.
2. The size of the two simulation cells representing the phases are equal and not changing during the course of the simulation. This, as was pointed out recently by Valleau, is quite important to obtain accurate values when extrapolating to infinite system size.
3. The speed of convergence both for the volumes and the number of particles can be arbitrarily fast as long as the integrator is stable.
4. The method is not connected to the insertion/deletion scheme, thus, any other suitable method can be used for direct chemical potential determination.

A disadvantage of our method with respect to the GEMC scheme is its non-Markovian character. However, this problem can be remedied by setting zeros to the higher order derivatives of the volume and number of particle related differential equations after the system moved to the neighborhood of the equilibrium. Then fluctuation amplitudes around the correct equilibrium values are identical of that of the GEMC scheme.

The method in its present form cannot be used for multicomponent fluid equilibria simulation with changing chemical composition. However, if the chemical constituents of the system are not very different, it is straightforward to generalize the algorithm with an additional equilibrium criterion and a subroutine which (when it is required) transforms one chemical species to another.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of OTKA Grant No. T24042 and the support of the NSF-MTA Collaborative Research Program through Grant No. INT-9603005.